

The likelihood for supernova neutrino analyses

A. Ianni,¹ G. Pagliaroli,^{1,2} A. Strumia,^{3,4} F. R. Torres,^{1,5} F. L. Villante,^{2,1} and F. Vissani¹

¹*INFN, Laboratori Nazionali del Gran Sasso, Assergi (AQ), Italy*

²*Dipartimento di Fisica, University dell'Aquila, Coppito (AQ), Italy*

³*Dipartimento di Fisica dell'Università di Pisa and INFN, Italy*

⁴*CERN, Theory division, CH-1211, Geneva 23, Switzerland*

⁵*Instituto de Física “Gleb Wataghin”, Universidade Estadual de Campinas, UNICAMP, Campinas, SP, Brasil*

We derive the event-by-event likelihood that allows to extract the complete information contained in the energy, time and direction of supernova neutrinos, and specify it in the case of SN1987A data. We resolve discrepancies in the previous literature, numerically relevant already in the concrete case of SN1987A data.

Preprint: LNGS/TH-01/09 and CERN-PH-TH/2009-115

I. INTRODUCTION

SN1987A neutrino events [1] prompted many dedicated analyses. Even if the number of detected neutrinos is limited, these analyses provide interesting limits on neutrino properties and clues on the core collapse mechanism.

The question of which likelihood should be adopted for supernova neutrino data analysis will become crucial after the next galactic supernova, when a much larger number of neutrino events will be collected. These events will carry information on neutrino properties mixed with information about supernova properties, so that we will need to jointly study their energy, time and direction distributions to try to extract all the relevant pieces of information. Therefore it will remain unpractical to bin the events and an event-by-event likelihood will remain the best tool for data analysis.

We present here the likelihood that should be ideally adopted for supernova neutrino data analysis. Our likelihood is more general than those already present in the literature [2, 3, 4]–[5]. Moreover, we resolve discrepancies in the previous literature, numerically relevant already in the concrete case of SN1987A data. We argue, in particular, that the analysis of SN1987A neutrino data by Lamb and Loredo [5] (LL), quoted since 2004 in the summary table of the Particle Data Group [6], uses a likelihood that incorrectly biases the analysis in favor of low energy events. We here present the correct likelihood, generalizing the ‘traditional’ form, advocated, e.g., by Jegerlehner, Neubig and Raffelt [4].

The structure of this paper is the following. In Sect. II we derive the general form of the likelihood. The application to a specific case of interest is discussed in Sect. III. Finally, in Sect. IV we compare our likelihood with other forms adopted for the analysis of SN1987A neutrinos, showing how the fitted parameters got biased.

II. DERIVATION OF THE LIKELIHOOD

A. General form of the likelihood

We write the expected event number in the i -th bin as:

$$n_i = dt_i d\mathbf{x}_i \frac{dN}{dt d\mathbf{x}}(t_i, \mathbf{x}_i), \quad (1)$$

where t_i represents the time coordinate, while \mathbf{x}_i indicates the set of all other observables (energy, position, direction, etc.) which define the properties of the i -bin. We suppose that the bin sizes $dt_i d\mathbf{x}_i$ are infinitesimally small so that the condition $n_i \ll 1$ holds true: therefore the probability that multiple events are collected in one bin is negligible and, thus, observing N_{ev} events corresponds to N_{ev} bins with 1 event, and all other bins with 0 events.

According to Poissonian statistics (see e.g., Appendix A of [7]) the associated likelihood is:

$$\mathcal{L} = \exp \left[- \sum_{j=1}^{N_{\text{bin}}} n_j \right] \times \prod_{i=1}^{N_{\text{ev}}} n_i, \quad (2)$$

where the sum in the exponent runs over all N_{bin} bins and gives the total number of expected events, while the product runs over all N_{ev} observed events. As usual, one can convert this into a χ^2 distribution as $\mathcal{L} = e^{-\chi^2/2}$.

B. Distinguishing between signal and background

Let us consider the case when the detected events are due to a signal S , reprocessed in the detector through a response function \mathcal{R} , and to a known (measured) background process B . We have:

$$\frac{dN}{dt d\mathbf{x}}(t, \mathbf{x}) = B(t, \mathbf{x}) + \int dt' d\mathbf{x}' S(t', \mathbf{x}') \mathcal{R}(t', \mathbf{x}', t, \mathbf{x}) \quad (3)$$

The second term in the r.h.s. takes into account that a signal produced at the time t' and with coordinates \mathbf{x}' , due to detector response, could be observed with a probability $\mathcal{R}(t', \mathbf{x}', t, \mathbf{x})$ at a different time t and coordinate \mathbf{x} .

By integrating over all possible detection times and coordinates, we introduce the general form of the detection efficiency:

$$\eta(t', \mathbf{x}') \equiv \int dt d\mathbf{x} \mathcal{R}(t', \mathbf{x}', t, \mathbf{x}). \quad (4)$$

The efficiency obeys the condition $0 \leq \eta \leq 1$, if we describe a situation when the events can be lost. By factoring out η we define the smearing (or error) function \mathcal{G}

$$\mathcal{G}(t', \mathbf{x}', t, \mathbf{x}) \equiv \mathcal{R}(t', \mathbf{x}', t, \mathbf{x}) / \eta(t', \mathbf{x}') \quad (5)$$

normalized to unity:

$$\int dt d\mathbf{x} \mathcal{G}(t', \mathbf{x}', t, \mathbf{x}) = 1. \quad (6)$$

The background B , the efficiency η and the smearing \mathcal{G} describe the experimental apparatus. Assuming that they are known, we can use an experimental result to learn on the signal S , by the study of the likelihood function of Eq. (2) together with (3).

C. Simplifications

In the case of interest, it is possible to further simplify the problem by relying on the following assumptions:

(i) We assume that the response function factorizes in the time and in the coordinates as follows

$$\mathcal{R}(t', \mathbf{x}', t, \mathbf{x}) = r(t', t) \mathcal{R}(\mathbf{x}', \mathbf{x}). \quad (7)$$

We introduce the time-independent efficiency in the observables $\eta(\mathbf{x}')$, defined in analogy to Eq. (4):

$$\eta(\mathbf{x}') \equiv \int d\mathbf{x} \mathcal{R}(\mathbf{x}', \mathbf{x}), \quad (8)$$

and the smearing function defined in analogy with Eq. (5):

$$\mathcal{G}(\mathbf{x}', \mathbf{x}) \equiv \mathcal{R}(\mathbf{x}', \mathbf{x}) / \eta(\mathbf{x}'). \quad (9)$$

Again, it is normalized to unity:

$$\int d\mathbf{x} \mathcal{G}(\mathbf{x}', \mathbf{x}) = 1. \quad (10)$$

We will discuss later the specific form of these expressions for SN1987A.

(ii) If the time t is measured with negligible error, we have

$$r(t', t) = \delta(t - t'), \quad (11)$$

possibly multiplied by a window function $w(t)$ to account for the dead time τ after an event, due to supernova or to background (for example, a muon), has been recorded. Concerning SN1987A data, only the *relative* time between events of the detectors was measured precisely; one

needs to take into account the uncertainty in the absolute time of the Kamiokande-II and Baksan events.

(iii) We can finally assume that the background does not depend on the time, namely

$$B(t, \mathbf{x}) = B(\mathbf{x}) \quad (12)$$

possibly, multiplied by $w(t)$ to take into account for the absence of any events, including those due to background, during dead time. Eq. (12) implies that the background can be *measured* in the period when the signal is absent (as for SN1987A).

With these assumptions, Eq. (3) simplifies to:

$$\frac{dN}{dt d\mathbf{x}}(t, \mathbf{x}) = B(\mathbf{x}) + \int d\mathbf{x}' \mathcal{G}(\mathbf{x}', \mathbf{x}) \eta(\mathbf{x}') S(t, \mathbf{x}'). \quad (13)$$

Then, assuming that the N_{ev} events \mathbf{x}_i have been measured at time t_i , the likelihood in Eq. (2) becomes:

$$\mathcal{L} = e^{-\int dt d\mathbf{x} B(\mathbf{x}) - \int dt d\mathbf{x}' \eta(\mathbf{x}') S(t, \mathbf{x}')} \times \prod_{i=1}^{N_{\text{ev}}} [B(\mathbf{x}_i) + \int d\mathbf{x}' \mathcal{G}(\mathbf{x}', \mathbf{x}_i) \eta(\mathbf{x}') S(t_i, \mathbf{x}')] dt_i d\mathbf{x}_i, \quad (14)$$

where, in the exponent, we replaced the sum over all infinitesimal bins with an integral and used (10). By dropping constant factors, that are irrelevant for estimating the parameters that control the theoretical expression of the signal rate S , and replacing \mathbf{x}' with \mathbf{x} , we get

$$\mathcal{L} = e^{-\int dt d\mathbf{x} \eta(\mathbf{x}) S(t, \mathbf{x})} \times \prod_{i=1}^{N_{\text{ev}}} [B(\mathbf{x}_i) + \int d\mathbf{x} \mathcal{G}(\mathbf{x}, \mathbf{x}_i) \eta(\mathbf{x}) S(t_i, \mathbf{x})]. \quad (15)$$

This form of the likelihood is general enough for the purpose of analyzing SN1987A neutrinos. Moreover, this is a generalization of the likelihood advocated in [8] for the study of radioactive decays, when the time of occurrence of each event is measured.

As we already discussed, the dead time can be taken into account by extending the time integral in the exponent only to the time when the detector is on, thereby removing the time intervals where data taking was stopped after each candidate signal event. As long as τ is small enough, one can equivalently take into account the dead time due to background events by multiplying the integrand in the exponent of (15) by the average live-time fraction, $1 - \tau b_\mu$, where b_μ is the time-averaged background event rate. Compare it with the discussion of [9], further elaborated in [5].

III. APPLICATION TO KAMIOKANDE-II

In order to specify the general formulæ, we choose a concrete and important example: we discuss the likelihood for the water Čerenkov detector Kamiokande-II.

A. Generalities

In this subsection, we collect some useful definitions.

The variables that characterize an event are:

$$\mathbf{x}_i = \{E_i \text{ (energy)}, \hat{n}_i \text{ (direction)}, \vec{r}_i \text{ (position)}\}. \quad (16)$$

For concreteness, we consider events resulting from the reaction $\bar{\nu}_e p \rightarrow ne^+$ when a positron is detected through its Čerenkov light; similar considerations apply to the elastic scattering reaction or the charged current reactions with nuclei.

In the construction of the likelihood 3 different directions are relevant: the direction \hat{n}_* of SN1987A; the *reconstructed* direction \hat{n}_i of each event; the *true* direction \hat{n} of the positrons produced by the detection process. For each event, the first 2 directions are fixed, while we have to integrate on the true direction of the positron, taking into account the detector response and the reconstructed event direction, as described in Eq. (15). To do this, it is convenient to use an “event-centric” system in which:

- 1) The reconstructed positron direction is along the z axis,

$$\hat{n}_i = (0, 0, 1). \quad (17)$$

- 2) The true positron direction is in the generic direction:

$$\hat{n} = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta), \quad (18)$$

so that

$$\cos \theta = \hat{n}_i \hat{n}. \quad (19)$$

Thus, θ is the opening angle around the reconstructed direction and φ is the azimuthal angle. The experimental collaborations usually quote the error on the angle $\delta\theta_i$ between the true and the reconstructed direction for each bin (rather than the error on the direction vector $\delta\hat{n}_i$ itself).

- 3) Finally, we have the vector \hat{n}_* pointing in the direction of the supernova. This, without losing in generality, can be chosen in the plane $x - z$:

$$\hat{n}_* = (\sin \theta_i, 0, \cos \theta_i), \quad (20)$$

so that

$$\cos \theta_i = \hat{n}_* \hat{n}_i. \quad (21)$$

B. Smearing function

In the simplest approximation, we can describe the smearing function by assuming that it factorizes according to:

$$\mathcal{G}(\mathbf{x}, \mathbf{x}_i) = G_1(E - E_i, \sigma_1) G_2(\hat{n} - \hat{n}_i, \sigma_2) G_3(\vec{r} - \vec{r}_i, \sigma_3) \quad (22)$$

where we denote by

$$G_n(\vec{x}, \sigma) = \frac{\exp(-\vec{x}^2/2\sigma^2)}{N_n(\sqrt{2\pi}\sigma)^n}, \quad (23)$$

a standard Gaussian in n dimensions. We include a normalization factor N_n to describe the presence of physical boundaries, like e.g., the fact that $\theta \in [0, \pi]$ and $\varphi \in [0, 2\pi]$ when we integrate over the possible directions of \hat{n} .

The quantities $\sigma_{1,2,3}$ are functions of the variables of Eq. (16). We identify $\sigma_{1,2,3}$ in the point $E = E_i, \hat{n} = \hat{n}_i, \vec{r} = \vec{r}_i$ with the error for the i -th event quoted by the experimental collaborations, e.g.,

$$\begin{aligned} \sigma_1(E_i, \hat{n}_i, \vec{r}_i) &= \delta E_i \\ \sigma_2(E_i, \hat{n}_i, \vec{r}_i) &= \delta n_i \\ \sigma_3(E_i, \hat{n}_i, \vec{r}_i) &= \delta r_i \end{aligned} \quad (24)$$

thus, we further approximate the smearing function:

$$\mathcal{G}(\mathbf{x}, \mathbf{x}_i) \approx G_1(E - E_i, \delta E_i) G_2(\hat{n} - \hat{n}_i, \delta n_i) G_3(\vec{r} - \vec{r}_i, \delta r_i) \quad (25)$$

in the vicinity of each point where the likelihood should be evaluated — see Eq. (15).¹

C. Remarks on the angular distribution

The expressions for energy and position smearing functions are essentially standard and do not need particular attention. The angular distribution requires, instead, a more detailed discussion.

First, we discuss the connection between the error δn_i to be inserted into Eq. (24) and the 1-sigma error on angle $\delta\theta_i$ indicated by the experimental collaborations. In the Gaussian assumption, the function G_2 can be written as:

$$G_2(\hat{n} - \hat{n}_i, \delta n_i) d\hat{n} = \frac{d\varphi}{2\pi} \times dc \frac{d\rho}{dc}(c, \delta n_i) \quad (26)$$

where the function $d\rho/dc$, given by:

$$\frac{d\rho}{dc} = \frac{1}{N_2 \delta n_i^2} \exp\left[-\frac{1-c}{\delta n_i^2}\right], \quad (27)$$

describes the distribution of the angle $c = \cos \theta = \hat{n}_i \hat{n}$ between the true and the reconstructed direction; the azimuthal angle φ is uniformly distributed as appropriate for an unbiased detector. The normalization factor N_2 can be explicitly calculated

$$N_2 = 1 - \exp(-2/\delta n_i^2), \quad (28)$$

¹ An (arguably) more refined approximation that takes into account the Poisson nature of the photoelectron detection can be obtained by multiplying the constant errors δE_i , δn_i and δr_i , by $\sqrt{E_i/E}$. We note, in this respect, that the normalization condition in Eq. (10) is obeyed even in the general case, when the functions $\sigma_{1,2,3}$ vary with the true coordinates of positron, since one integrates over the *reconstructed* coordinates.

and it is very close to one for the typical case $\delta n_i \ll 1$. We calculate δn_i by requiring that:

$$\int_0^{\delta\theta_i} d\theta \frac{d\rho}{d\theta}(\theta, \delta n_i) = 0.683, \quad (29)$$

for the 1 sigma error $\delta\theta_i$ corresponds to the ≈ 0.683 confidence level. For small δn_i , we get easily:

$$\delta n_i \simeq 0.660 \delta\theta_i \times (1 - \delta\theta_i^2/24). \quad (30)$$

Typically the first term provides an adequate approximation for the quantity we search, δn_i .

Now, we discuss a possible improvement of the Gaussian assumption for the distribution G_2 . Experimental investigations of the Super-Kamiokande collaboration [11] have shown that the tails of the angular distribution fall slower than $\exp(-\text{cte} \cdot \theta^2)$ and resemble more closely $\exp(-\text{cte} \cdot \theta)$; compare also with App. C of [12]. This suggests to release the Gaussian approximation for the distribution on the directions and to replace $G_2 \rightarrow \exp(-|\hat{n} - \hat{n}_i|/\delta n_i)$. Thus, the distribution over the cosine becomes:

$$\frac{d\rho}{dc} = \frac{1}{N_2 \delta n_i^2} \exp\left[-\frac{\sqrt{2(1-c)}}{\delta n_i}\right], \quad (31)$$

where the proportionality constant

$$N_2 = 1 - (1 + 2/\delta n_i) \exp(-2/\delta n_i), \quad (32)$$

is again close to one for small δn_i . By imposing the condition (29) and considering again the limit of small δn_i we calculate the new expression for δn_i , obtaining:

$$\delta n_i \simeq 0.424 \delta\theta_i \times (1 - \delta\theta_i^2/24). \quad (33)$$

where, as in Eq. (30), the first term is typically sufficient.

The two distributions are depicted in Fig. 1 for a specific value of $\delta\theta_i$. It is worthwhile to note various features of Eqs. (27) and (31):

1. When considered as functions of the direction we see that they both depend only $|\hat{n} - \hat{n}_i|$ and have a maximum for $\hat{n} = \hat{n}_i$, as it should be.
2. It is easy to treat them analytically, which is a welcome property to use them in a likelihood.
3. It is simple to study their limit for small δn_i by replacing $\sin\theta \rightarrow \theta$ and $\cos\theta \rightarrow 1 - \theta^2/2$, which makes their analytical treatment even simpler.²

² When only the second replacement is done, Eq. (27) coincides with the form commonly used in the literature [10] and Eq. (31) practically coincides with the form given in [12]. In fact, the exponential term in Eq. C1 of [12] can be neglected in comparison to the linear term x for all relevant energies.

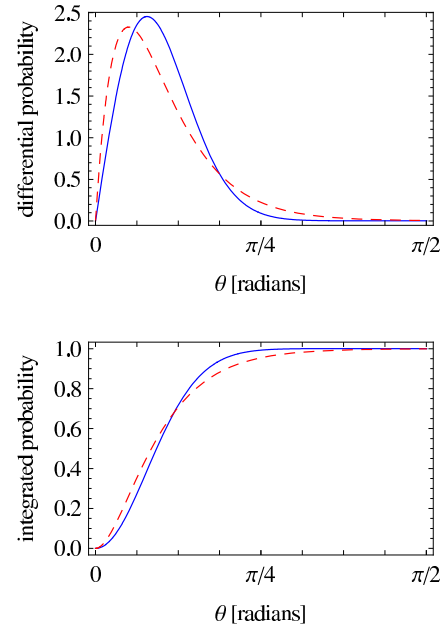


FIG. 1: Angular distribution (top) and related cumulative distribution (bottom) as a function of the angle θ . The continuous curves show Eq. (27) and the dashed curves show Eq. (31). We here assumed $\delta\theta = 83^\circ \sqrt{\text{MeV}/E}$, as appropriate for Super-Kamiokande [11] and a typical energy of an elastic scattering event, $E = 15$ MeV, so that $\delta\theta = 0.374$.

4. For a fixed $\delta\theta_i$, we see that δn_i is smaller in the second case; thus, the maximum at $\cos\theta = 0$ is higher in the second case.
5. The most probable angle is $\theta \simeq \delta n_i$ in both cases; thus it is smaller in the second case.

A choice between Eq. (27) and Eq. (31) (or other reasonable approximations) is not critical for the analysis of SN1987A in view of the limited event sample. However, the use of an appropriate distribution is potentially important for the analysis of the elastic scattering events from a future supernova in a water Čerenkov detector.

D. From the idealized to the actual likelihood

We are now in the position to provide a concrete and useful expression for the likelihood that takes into account the reported information on the data and on the detector response.

We recall that:

- 1) The signal is expected to be uniformly distributed inside the detector. The angular dependence of the signal arises from the angular distribution of positrons produced by $\bar{\nu}_e p \rightarrow n e^+$. This can be expressed as a function of the angle $\hat{n}\hat{n}_*$ between the direction \hat{n} of the produced positron and the direction \hat{n}_* of the SN1987A. We can

thus replace in Eq. (15):

$$S(t, \mathbf{x}) \rightarrow \frac{S(t, E, \hat{n}\hat{n}_*)}{2\pi V} \quad (34)$$

where V is the volume of the detector, the product $\hat{n}\hat{n}_*$ can be expressed through Eqs. (18) and (20) as:

$$\hat{n} \cdot \hat{n}_* = \cos \theta_i \cos \theta + \sin \theta_i \sin \theta \cos \varphi, \quad (35)$$

and the factor 2π accounts for the fact that positron directions are uniformly distributed with respect to rotations around \hat{n}_* .

2) We assume that the background does not depend on time, direction and position. We indicate the total background counting rate as a function of the energy with $\bar{B}(E)$ and we replace in Eq. (15):

$$B(\mathbf{x}_i) \rightarrow \frac{\bar{B}(E)}{4\pi V}; \quad (36)$$

3) The average efficiency of the detector as a function of the energy $\bar{\eta}(E)$ is known. We assume that the efficiency does not depend on time, position and direction, so that we can replace in Eq. (15):

$$\eta(E, \hat{n}, \vec{r}) \rightarrow \bar{\eta}(E); \quad (37)$$

4) The errors on the energy δE_i and on the direction $\delta \theta_i$ in the neighbourhood of the given datum are known. We additionally indicate by δr_i the value of the error on the position, on which we have only limited information, $\delta r_i \sim 1$ m at 10 MeV.

At this point, we have all the elements to write the concrete form of the likelihood. Integrating away the Gaussian on the positions G_3 and omitting the constant factor $1/(2\pi V)^{N_{\text{ev}}}$ we get from Eq. (15):

$$\mathcal{L} = e^{-\int_T dt \int dE \int dc \bar{\eta}(E) S(t, E, c)} \prod_{i=1}^{N_{\text{ev}}} \left[\frac{\bar{B}(E_i)}{2} + \int \bar{\eta}(E) dE G_1(E - E_i, \delta E_i) \int \frac{d\varphi}{2\pi} \int dc \frac{d\rho}{dc}(c, \delta n_i) \times S(t_i, E, c_i c + s_i s c_\varphi) \right], \quad (38)$$

where Eq. (35) has been rewritten using the intuitive shorthands $\cos \theta \rightarrow c$, $\sin \theta_i \rightarrow s_i$, etc.

IV. COMPARISON WITH THE LITERATURE

Here, we compare our likelihood, Eq. (38), with certain other likelihoods present in the recent literature and currently used for the analysis of SN1987A events.

A. Jegerlehner, Neubig and Raffelt [4]

The first likelihood is Eq. (15) of [4]:

$$\mathcal{L}^{JNR} = C e^{-\int_0^\infty n(E) dE} \prod_{i=1}^{N_{\text{obs}}} n(E_i). \quad (39)$$

This is an approximation of our likelihood, in that the background has been neglected and the time and angular distribution are integrated (averaged) over; in other words, only the energy distribution is considered. However, this expression is in direct correspondence with Eq. (38), when $B \rightarrow 0$ and $\delta \theta_i \rightarrow 0$, and if we take as the definition of $n(E_i)$ the one given in Eqs. (18,19,21) of [4]. Furthermore, the expression of [4] agrees with Eq. (2).

B. Lamb and Loredo [5]

The other likelihood that we consider is the one advocated by Lamb and Loredo [5]. This is given by their Eq. (3.18) which, rewritten in our notations, reads

$$\mathcal{L}^{LL} = e^{-\int_T dt dE \bar{\eta}(E) S(E, t)} \times \prod_{i=1}^{N_{\text{ev}}} [\int dE \mathcal{L}_i(E) S(E, t_i) + \bar{B}(E_i)], \quad (40)$$

where we neglected dead-time, as appropriate for Kamiokande-II, and dropped the information about the angular distribution. Quoting [5]: *Our derivation of the likelihood function reveals errors in previous attempts to account for the energy dependence of the efficiencies of the neutrino detectors; we show that these errors significantly corrupt previous inferences.* Indeed, the likelihood advocated by LL has been shown to have an important impact for the analysis of data also by similar and independent analyses of SN1987A observations [13].

We would like, however, to draw the discussion on the correctness of the likelihood of LL. We see that the LL expression, Eq. (40), coincides with our Eq. (38) *only if we identify* the function \mathcal{L}_i with the energy response function of the detector:

$$\mathcal{L}_i(E) \stackrel{(?)}{=} G_1(E - E_i, \sigma_i) \bar{\eta}(E), \quad (41)$$

where G_1 is the Gaussian smearing of Eq. (23).

This is not the case for LL who instead claim (see their Eq. (3.21)):

$$\mathcal{L}_i(E) \stackrel{(!)}{=} G_1(E - E_i, \sigma_i) \Theta(E - E_0), \quad (42)$$

where E_0 is assumed to be the maximum energy where the efficiency vanishes (i.e., the minimum detectable energy) and Θ is the step function.

The only special case in which the LL likelihood coincides with our result is when the average efficiency is assumed to be a step function $\bar{\eta}(E) = \Theta(E - E_0)$. In general, this is not the case and the efficiency is a continuously growing function of the energy. The LL likelihood therefore incorrectly biases the analysis in favor of low energy events. The quantitative effect of this bias on data analysis will be discussed further in Sect. IV E.

The above remarks amount to the consideration that the likelihood of Lamb and Loredo does *not* follow from the formal construction described in Sects. II and III.

However, it is instructive to point out more directly the profound principle problem of the LL likelihood.

We begin noting that Eq. (40) has been derived by omitting constant terms from

$$\mathcal{P}^{LL} = e^{-\int_T dt dE [\bar{\eta}(E) S(E, t) + \bar{B}(E)]} \times \prod_{i=1}^{N_{\text{ev}}} [\int dE \mathcal{L}_i(E) S(E, t_i) + \bar{B}(E_i)] dt_i dE_i, \quad (43)$$

which should represent the *probability* that a given experimental result is obtained. This expression is supposed to have a general validity. Then consider a simple limiting case: only one bin, with dimension $\Delta t \times \Delta E$ and with energy above E_0 ; no background, $B(E) \equiv 0$; a constant signal, $S(E, t) \equiv S$; a constant efficiency, $\bar{\eta}(E) \equiv \eta$; a perfect energy resolution, $\delta E_i \rightarrow 0$; a very small expected number of events, $n \equiv \eta S \Delta t \Delta E \ll 1$. In these assumptions, the most probable outcome is the case $N_{\text{ev}} = 0$, followed by the case $N_{\text{ev}} = 1$; the probability of other possible results is negligible. From Eq. (43), we calculate the probabilities of the cases when no event and one event are observed: $\mathcal{P}_0^{LL} = 1 - n$ and $\mathcal{P}_1^{LL} = n/\eta$ respectively. Their sum violates the basic principle according to which the sum of the probability for all possible results should be equal to one.

C. Loredo [14] and Bernstein et al. [15]

We comment here on a likelihood that was *not* proposed for the analysis of supernova neutrinos, but that it is strictly connected with the previous one.

The same position as in Eq. (42) was made in [14], where Loredo defines the quantity $\ell_i(m) = p(d_i|m)$. This quantity, that evidently corresponds to the quantity \mathcal{L}_i discussed above, is claimed to be independent on the detection efficiency. This position caused the criticism of Loredo [14] to the likelihood advocated for the analysis of trans-neptunian objects of Bernstein *et al.* 2004 [15] (see Eq. (A8) of [15]). Again, we find the position of [14] unjustified, while we agree with [15]. In particular, Eq. (A4) of [15] expresses the statement that the response function *does* contain the detection efficiency; it corresponds strictly to our Eq. (4).

D. Pagliaroli et al. [13]

Finally, in the analyses performed by some of us [13], Eq. (38) was used, simplified to the case $\delta n_i \rightarrow 0$ to take into account the mild angular dependence of the $\bar{\nu}_e p \rightarrow n e^+$ reaction. This simplification does not affect significantly the analysis of SN1987A events.

E. Numerical comparison of the likelihoods

The bias implied by the likelihood of ref. [5] is numerically important already for the analysis of SN1987A events, as found in [5] and confirmed by [13].

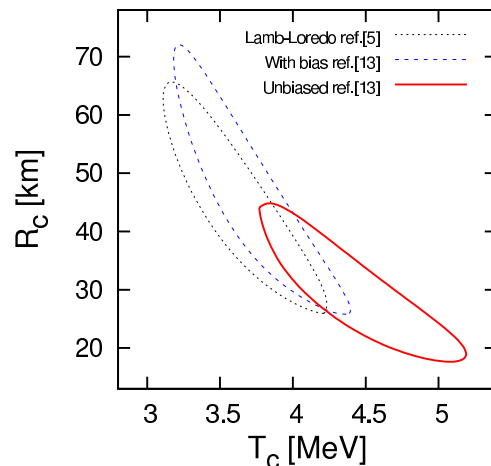


FIG. 2: Comparison of the joint 68% C.L. (2 dof) regions for the cooling radius R_c and the initial temperature T_c obtained from two analyses of SN1987A data based on the exponential cooling model. When Eq. (42) is adopted, there is a good agreement between the results of [5] (dotted line) and those of [13] (dashed line); the small discrepancies can be ascribed to the different statistical procedures (Bayesian in [5] and frequentist in [13]) and to a different numerical treatment of the data. The effect of switching from Eq. (42) (dashed line) to Eq. (41) (continuous line) is much more significant [13].

In order to illustrate this point better, we recall certain results obtained in the previous analyses. Let us begin by considering the conventional exponential cooling model, in which the $\bar{\nu}_e$ temperature decreases exponentially with the time and the neutrino-radius R_c is constant. As evident from Fig. 2, the use of Eq. (42) rather than Eq. (41) leads to important differences on the inferred values of the parameters. We note in particular that the (well-known) difference between R_c and the expected size of the neutron star radius, $R_{\text{ns}} \sim 15$ km, is amplified when we adopt Eq. (42) (i.e., when we bias the analysis). This outcome can be easily understood: the bias in favor of low energy events implies that T_c (that is proportional to the average energy of the electron antineutrinos) will decrease; thus, R_c has to increase to keep the number of events constant.

The bias will be even more important for the analysis of a future galactic supernova, since the number of collected events will be much larger and the errors on the parameters are expected to scale as the square root of the number of the events.

Indeed, the analysis of the 29 events collected by Kamiokande-II, IMB and Baksan in an extension of the exponential cooling model leads to $R_c = 16^{+9}_{-5}$ km and $T_c = 4.6^{+0.7}_{-0.6}$ MeV [13]. A recent analysis of simulated events from a future supernova, that assumed the same antineutrino emission model (that includes an initial phase of intense emission), the same central values as found from SN1987A (in particular, $R_c = 16$ km and $T_c = 4.6$ MeV), and a supernova located at a distance of 20 kpc (i.e., a data set 30 times larger), leads to the con-

clusions that the parameters are correctly reconstructed when we use Eq. (41). Moreover, when we combine the results of the simulations, we can estimate the average values of the parameters and of their expected errors: $R_c = 15.4 \pm 0.9$ km and $T_c = 4.6 \pm 0.1$ MeV [16].

The comparison with the values from SN1987A reveals that the errors are expected to decrease by about six times, which is similar to the improvement that we can ascribe to the increased number of data. We are lead to the conclusion that, after a future galactic supernova, the allowed regions in Fig. 2 should shrink by a similar factor in linear scale, making the effect of the bias much more important.

V. SUMMARY

We constructed the general likelihood for supernova data analysis, Eq. (15), and specified it to the analysis of SN1987A, Eq. (38). We have compared this likelihood

with other forms advocated in the scientific literature. While our likelihood is a generalization of the likelihoods traditionally adopted for the analysis of SN1987A events (or in general for the study of rare processes), it is in disagreement with other ones. Reasons and consequences of these disagreements are discussed.

Acknowledgments

We thank A. Dighe, E. Lisi, D. Montanino and F. Teranova for useful discussions. This work was partly supported by High Energy Astrophysics Studies contract number ASI-INAF I/088/06/0; MIUR grant for the Projects of National Interest PRIN 2006 “Astroparticle Physics”; European FP6 Network “UniverseNet” MRTN-CT-2006-035863. F.R.T. thanks CAPES (grant number 3247-08-2) for financial support and INFN Gran Sasso for hospitality.

-
- [1] R.M. Bionta *et al.*, Phys. Rev. Lett. 58 (1987) 1494 and C.B. Bratton *et al.*, Phys. Rev. D 37 (1988) 3361; K. Hirata *et al.*, Phys. Rev. Lett. 58 (1987) 1490; K.S. Hirata *et al.*, Phys. Rev. D 38 (1988) 448; E.N. Alekseev, L.N. Alekseeva, I.V. Krivosheina and V.I. Volchenko, Phys.Lett.B 205 (1988) 209.
 - [2] D.N. Spergel, T. Piran, A. Loeb, J. Goodman and J.N. Bahcall, Science 237 (1987) 1471.
 - [3] T.J. Loredo, D.Q. Lamb, Annals N. Y. Acad. Sci. 571 (1989) 601.
 - [4] B. Jegerlehner, F. Neubig and G. Raffelt, Phys. Rev. D 54, 1194 (1996).
 - [5] T.J. Loredo and D.Q. Lamb, Phys. Rev. D 65 (2002) 063002.
 - [6] Summary table on “ $\bar{\nu}$ mass” of Particle Data Group: S. Eidelman *et al.* Phys. Lett. B 592 (2004), page 439; W.M. Yao *et al.*, J. Phys. G 33 (2006), page 472; C. Am- sler *et al.*, Phys. Lett. B 667 (2008), page 517.
 - [7] A. Strumia and F. Vissani, JHEP 0111 (2001) 048.
 - [8] B.T. Cleveland, Nucl. Instrum. Meth. 214 (1983) 451.
 - [9] P.J. Kernan, L.M. Krauss, Nucl. Phys. B 437 (1995) 243.
 - [10] D. Kielczewska, Phys. Rev. D 41, 2967 (1990).
 - [11] M. Nakahata *et al.*, Nucl. Instr. Meth. A421 (1999) 113.
 - [12] R. Tomas, D. Semikoz, G.G. Raffelt, M. Kachelriess and A.S. Dighe, Phys. Rev. D 68 (2003) 093013.
 - [13] G. Pagliaroli, F. Vissani, M.L. Costantini and A. Ianni, Astropart. Phys. 31 (2009) 163.
 - [14] T.J. Loredo, AIP Conf. Proc. 735 (2004) 195.
 - [15] G.M. Bernstein *et al.*, Astronomical J. 128 (2004) 1364.
 - [16] G. Pagliaroli, F. Vissani, E. Coccia and W. Fulgione, arXiv:0903.1191 [hep-ph], accepted for publication in PRL.